

520.42834X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): TAKAMURA, et al.
Serial No.: Not assigned
Filed: July 30, 2003
Title: COMPUTER SYSTEM HAVING A VIRTUALIZED I/O DEVICE
Group: Not assigned

LETTER CLAIMING RIGHT OF PRIORITY

Mail Stop Patent Application
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

July 30, 2003

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on Japanese Application No.(s) 2003-046885 filed February 25, 2003.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



Ronald J. Shore
Registration No. 28,577

RJS/amr
Attachment
(703) 312-6600

日 本 国 特 許 庁

JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

2003年 2月25日

出 願 番 号

Application Number:

特願2003-046885

[ST.10/C]:

[JP2003-046885]

出 願 人

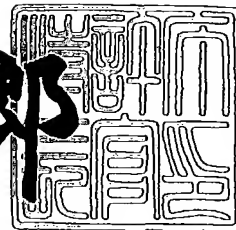
Applicant(s):

株式会社日立製作所

2003年 5月 6日

特許庁長官
Commissioner,
Japan Patent Office

太田信一郎



出証番号 出証特2003-3032211

【書類名】 特許願

【整理番号】 NT03P0117

【提出日】 平成15年 2月25日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 13/10

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内

【氏名】 高村 明裕

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内

【氏名】 三木 良雄

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100068504

【弁理士】

【氏名又は名称】 小川 勝男

【電話番号】 03-3661-0071

【選任した代理人】

【識別番号】 100086656

【弁理士】

【氏名又は名称】 田中 恭助

【電話番号】 03-3661-0071

【選任した代理人】

【識別番号】 100094352

【弁理士】

【氏名又は名称】 佐々木 孝

【電話番号】 03-3661-0071

【手数料の表示】

【予納台帳番号】 081423

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】明細書

【発明の名称】仮想化された I/O デバイスをもつ計算機システム

【特許請求の範囲】

【請求項 1】

第 1 のハイパバイザと、その上で動作する第 1 の OS とを有するクライアントコンピュータと、第 2 のハイパバイザと、その上で動作する第 2 の OS と、物理 I/O デバイスとを有するサーバコンピュータとを備え、前記クライアントコンピュータと前記サーバコンピュータとはネットワークで接続されており、前記第 1 のハイパバイザは前記サーバコンピュータの前記物理 I/O デバイスへのアクセスである論理 I/O デバイスアクセスを検出する論理 I/O デバイスアクセス検出手段と、前記論理 I/O デバイスアクセスがあったとき、前記論理 I/O デバイスアクセスのコマンドを前記ネットワークを介して前記サーバコンピュータへ送信する仮想 I/O クライアント処理とを有し、前記第 2 のハイパバイザは前記コマンドを前記ネットワークを介して受信して前記物理 I/O デバイスへコマンドを発行する仮想 I/O サーバ処理を有することを特徴とする仮想化された I/O デバイスをもつ計算機システム。

【請求項 2】

前記クライアントコンピュータと前記サーバコンピュータの間での通信は前記第 1 と第 2 のハイパバイザが定めるプロトコルで行なわれることを特徴とする請求項 1 記載の仮想化された I/O デバイスをもつ計算機システム。

【請求項 3】

前記クライアントコンピュータと前記サーバコンピュータの間での通信は前記サーバコンピュータ上で動作する第 2 の OS のプロトコルで行なわれることを特徴とする請求項 1 記載の仮想化された I/O デバイスをもつ計算機システム。

【請求項 4】

前記仮想 I/O クライアント処理は前記論理 I/O デバイスアクセスのコマンドを前記サーバコンピュータの第 2 の OS が解釈可能なプロトコルに変換してから前記コマンドを前記サーバコンピュータに送ることを特徴とする請求項 1 記載の仮想化された I/O デバイスをもつ計算機システム。

【請求項5】

前記クライアントコンピュータは記憶保護制御機能を有し、更に、前記記憶保護制御機能を用いて前記論理 I / O デバイスへの読み出しコマンドと書き込みコマンドとを検出し前記第 1 のハイパバイザへ制御を移す手段を有することを特徴とする請求項 1 記載の仮想化された I / O デバイスをもつ計算機システム。

【請求項6】

前記クライアントコンピュータにおいて、特定のメモリアドレスに対する読み出し/書き込みが起きたときに記憶保護割込が発生し、記憶保護割込となった原因が前記論理 I / O デバイスへの読み出しコマンドの発行もしくは書き込みコマンドの発行であれば前記仮想 I / O クライアント処理を呼び出す手段を有することを特徴とする請求項 5 記載の仮想化された I / O デバイスをもつ計算機システム。

【請求項7】

前記第 1 の OS はアプリケーションプログラムが指定したファイルを格納している I / O デバイスに対応するデバイスドライバを検索し、この検索によって見つかったデバイスドライバを呼び出す手段を有し、前記デバイスドライバは読み出しコマンドを発行し、読み出しコマンドを論理 I / O デバイスのメモリアドレスに書き込みを行なうことを特徴とする請求項 6 記載の仮想化された I / O デバイスをもつ計算機システム。

【請求項8】

前記仮想 I / O クライアント処理は I / O デバイスの制御用メモリアドレスに書き込まれたデータが読み出しコマンドか書き込みコマンドかを判定する手段と、読み出しコマンドである場合は読み出しコマンドと読み出しコマンドのパラメータを前記サーバコンピュータへ送信する手段と、前記サーバコンピュータが I / O 読み出し処理で得たデータを受信する手段と、書き込みコマンドである場合は書き込みコマンドと書き込みコマンドのパラメータを前記サーバコンピュータに送信する手段と、書き込みデータを前記サーバコンピュータに送信する手段とを有することを特徴とする請求項 1 記載の仮想化された I / O デバイスをもつ計算機システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ネットワークを用いて接続された計算ノードをもつコンピュータシステムにおいて、特定の計算ノードに接続されたI/Oデバイスを他の計算ノードからでもアクセス可能なコンピュータシステムに関する。

【0002】

【従来の技術】

コンピュータシステムは、ハードウェアコストの削減、システム構築の容易化などの理由から、機能の異なる複数の計算ノードをネットワークで結合して構築することが多い。しかしながら、この様に機能の異なる複数の計算ノードがネットワークで結合されたコンピュータシステムでは、特定の計算ノードにのみ接続されたI/Oデバイスがあることが多い。この時、他の計算ノードからこの特定のI/Oデバイスを使用するためには、I/Oデバイスが接続されている計算ノードにlog inしたり、ノード間のデータ転送が必要になる。

【0003】

このようなコンピュータシステムの利便性を高めるため、従来は、OS(オペレーティングシステム)が、特定の計算ノードに接続されたI/Oデバイスを任意の計算ノードからアクセス可能とする機能を提供していた。この機能をここではI/Oデバイスの仮想化機能と呼ぶ。以下、I/Oデバイスを持つ計算ノードをサーバコンピュータ、サーバコンピュータのI/Oデバイスにアクセスするコンピュータをクライアントコンピュータと呼ぶ。例えば、特開平10-21203号公報(特許文献1)に記載のように、クライアントコンピュータのOSがサーバコンピュータのI/O装置へのアクセス要求を出すために、クライアントコンピュータのOSがサーバコンピュータへのI/Oアクセスが出来る機能を持つOSでなくてはならない。

【0004】

OSによるI/Oデバイスの仮想化機能は以下の様に実現されている。クライアントコンピュータにあるアプリケーションプログラムが、サーバコンピュータに接

続されているI/Oデバイス上にあるファイルを読み出す場合を考える。クライアントコンピュータとサーバコンピュータはネットワークを介して接続されている。クライアントコンピュータにあるアプリケーションプログラムがファイルを読み出す場合、ファイルの読み出し要求はOSの一部であるファイルシステムに伝えられる。ファイルシステムでは、ファイルの論理名から、そのファイルを含むI/Oデバイスが接続されているサーバコンピュータとサーバコンピュータ上での論理ファイル名を得る。そして、クライアントコンピュータのファイルシステムは、ネットワークを介してサーバコンピュータのファイルシステムに対して、サーバコンピュータ上でのファイルの論理名を伝える。サーバコンピュータのOSの一部であるファイルシステムは、ファイルの論理名からそのファイルを含むI/OデバイスとI/Oデバイスへの入出力を行なうデバイスドライバを特定する。デバイスドライバは、I/Oデバイスに読み出しコマンドを発行してファイルを読み出す。サーバコンピュータのファイルシステムは読み出したファイルをネットワークを介してクライアントコンピュータのファイルシステムに送る。クライアントコンピュータのファイルシステムは、アプリケーションプログラムに読み出したファイルを送る。これらの一連の処理により、クライアントコンピュータにあるプログラムが、サーバコンピュータに接続されているI/Oデバイスにあるファイルを読み出すことができる。

【0005】

しかしながら、実際のコンピュータシステムでは、OSによるI/Oデバイスの仮想化機能は使用できないことが多い。これは、I/Oの仮想化機能は、多くの場合同一OS間でのみ使用可能であることと、通常、複数種類のOSが一つのコンピュータシステム内に混在することが原因である。I/Oの仮想化が同一OS間に限られるのは、OSごとにI/Oデバイスの仮想化方法が異なるからである。また、コンピュータシステムが複数種類のOSで構成されるのは、コンピュータシステムは一度に構築されるのではなく、拡張を繰り返して構築されることが多いが、拡張の際に、その時で最も条件のよいハードウェア、OS、ソフトウェアを選択するため、コンピュータシステム全体としては複数種類のOSが混在することになってしまうためである。

【 0 0 0 6 】

【特許文献 1】

特開平 1 0 - 2 1 2 0 3 号公報

【 0 0 0 7 】

【発明が解決しようとする課題】

上記従来技術では、ネットワークを用いて接続されたサーバコンピュータとクライアントコンピュータからなるコンピュータシステムにおいて、サーバコンピュータとクライアントコンピュータのOSが異なる場合に、サーバコンピュータに接続されたI/Oデバイスがクライアントコンピュータでは使用できないという問題がある。

【 0 0 0 8 】

本発明の目的は、ネットワークを用いて接続されたサーバコンピュータとクライアントコンピュータからなるコンピュータシステムにおいて、サーバコンピュータとクライアントコンピュータのOSが異なる場合でも、サーバコンピュータのOSもクライアントコンピュータのOSも変更することなく、サーバコンピュータに接続されたI/Oデバイスがクライアントコンピュータで使用することを可能にすることである。

【 0 0 0 9 】

【課題を解決するための手段】

上記目的を達成するため、本発明では、OSよりハードウェアに近い階層で動作するハイパバイザを設ける。クライアントコンピュータのハイパバイザには、オペレーティングシステムとオペレーティングシステム上のプログラムを変更することなく、サーバコンピュータのI/Oデバイスへのアクセスを検出するための論理I/Oデバイスアクセス検出部と、サーバコンピュータのI/Oデバイスへのアクセスがあった時に、サーバコンピュータのI/Oデバイスへのコマンドをネットワークを介して送信する仮想I/Oクライアント処理を有する。サーバコンピュータのハイパバイザには、I/Oデバイスへのコマンドをネットワークから受信して、I/Oデバイスへコマンド発行する仮想I/Oサーバ処理とを有する。

【 0 0 1 0 】

【発明の実施の形態】

本発明の第一の実施の形態では、サーバコンピュータとクライアントコンピュータの間での通信は、両者のハイパバイザが定めるプロトコルで通信が行なわれるものである。一方、第二の実施の形態では、サーバコンピュータとクライアントコンピュータの間での通信は、サーバコンピュータ上で動作するOSのプロトコルで行なわれる。

【0011】

図1に第一、第二の実施形態で共通に用いられるコンピュータシステムを示す。

本実施形態のコンピュータシステムは、1台以上のクライアントコンピュータ101と1台以上のサーバコンピュータ102と、これらのコンピュータを相互に結合するネットワーク103からなる。

【0012】

クライアントコンピュータ101は、プログラムとデータを格納するメモリ901とメモリ901内のプログラムに従って処理を行うプロセッサ900とネットワークに接続するためのNIA(ネットワークインターフェースアダプタ)902とプログラムとデータを保存するHDD(ハードディスク)903を有する。

【0013】

同様に、サーバコンピュータ102は、プログラムとデータを格納するメモリ911とメモリ911内のプログラムに従って処理を行うプロセッサ910とネットワークに接続するためのNIA912とプログラムとデータを保存するHDD913とI/Oデバイス914を有する。I/Oデバイス914は例えばリムーバブルディスクドライブであるとする。

【0014】

図2に、第一と第二の実施形態におけるHDD903の内容を示す。HDD903には、アプリケーションプログラム121とオペレーティングシステム122とハイパバイザ123とブートローダ124が格納されている。

【0015】

アプリケーションプログラム121 はファイルの読み出し210とファイルの書き

込み360を含むプログラムであり、ハイパバイザの機能によってサーバコンピュータ102に接続されたI/O デバイス914からの読み出しと書き込みを行なう。ハイパバイザ123には記憶保護割込処理300 と仮想I/Oクライアント処理260が含まれる。記憶保護割込処理300はアプリケーションプログラム121がファイルの読み出し210 とファイルの書き込み360を行なった時に、プロセッサ900の記憶保護制御を用いて論理I/Oデバイスへの読み出しコマンドと書き込みコマンドを検出し、ハイパバイザへ制御を移す。仮想I/Oクライアント処理260は記憶保護割込処理300 によって起動され、ネットワーク103を介してサーバコンピュータ102に接続されたI/O デバイス914からの読み出しと書き込みを行なう。ブートローダ124はクライアントコンピュータ101の起動時にオペレーティングシステム122の起動前にハイパバイザ123を起動させ、記憶保護割込が起きた時にハイパバイザ123の記憶保護割込処理300が呼び出されるようにする。

【 0 0 1 6 】

なお、本実施形態におけるアプリケーションプログラム121にはファイルの読み出し210とファイルの書き込み360が含まれているが、ファイルの読み出し210、ファイルの書き込み360のいずれか一方だけでも本実施形態と同様である。

【 0 0 1 7 】

図3に、第一の実施形態におけるHDD913の内容を示す。HDD913には、オペレーティングシステム132とハイパバイザ133とブートローダ134 が格納されている。

【 0 0 1 8 】

ハイパバイザ133には記憶保護割込処理380とNIA受信割込処理280とが含まれる。記憶保護割込処理380はオペレーティングシステム132のNIA受信割込処理が呼び出される前にハイパバイザ133のNIA受信割込処理280が呼び出されるようにする。NIA受信割込処理280では、クライアントコンピュータ101からネットワーク103を介してI/Oデバイス914の読み出し/書き込み要求が来た時に、仮想I/Oクライアント処理260を呼び出す。仮想I/Oクライアント処理260は読み出し/書き込み要求に応じてI/Oデバイスの読み出し/書き込みを行なう。ブートローダ134はサーバコンピュータ102の起動時にオペレーティングシステム122の起動前にハイパバイザ123を起動させ、記憶保護割込とNIA受信割込が起きた時に、それぞれハイ

ハイバイザ133 の記憶保護割込処理380とNIA受信割込処理280とが呼び出されるようにする。

【0019】

ブートローダ124はサーバコンピュータ102の起動時にオペレーティングシステム132の起動前にハイバイザ133を起動させ、NIA受信割込が起きた時にハイバイザ133のNIA受信割込処理280が呼び出されるようにする。

【0020】

図4に第一、第二の実施形態で共通に用いられるファイルの読み出し手順の詳細を示す。

【0021】

クライアントコンピュータ101のアプリケーションプログラム121がファイルの読み出しを行なう場合、まず、オペレーティングシステム122はデバイスドライバの検索211を実行する。デバイスドライバは、オペレーティングシステムでI/Oデバイスを使用可能にするためのソフトウェアである。デバイスドライバはI/Oデバイスごとに異なるため、プログラムが指定したファイルを格納しているI/Oデバイスに対応するデバイスドライバを検索する。

【0022】

次に、デバイスドライバの呼び出し212を実行して、デバイスドライバの検索211でみつかったデバイスドライバを呼び出す。さらに、デバイスドライバは、読み出しコマンド発行213を行ない、読み出しコマンドを論理I/Oデバイスの制御用メモリアドレスに書き込みを行なう。読み出しコマンド発行213により、プロセッサの記憶保護機能によって記憶保護割込が発生し、論理I/Oデバイスアクセス検出214が行なわれる。記憶保護割込によって制御がハイバイザの記憶保護割込処理300に移る。ハイバイザによって仮想I/Oクライアント処理260が行われ、サーバコンピュータのI/Oデバイスからの読み出しが完了する。

【0023】

図5に第一、第二の実施形態で共通に用いられるファイルの書き込み手順の詳細を示す。書き込みも図4のファイルの読み出し手順と基本的には同一であるが、読み出しコマンド発行213の代わりに書き込みコマンドの発行363を行なうこと

が異なる。

【 0 0 2 4 】

図6に第一、第二の実施形態で共通に用いられる記憶保護割込処理300の手順の詳細を示す。記憶保護割込処理300はクライアントコンピュータ101で実行されるハイパバイザ内のプログラムで、特定のメモリアドレスに対する読み出し/書き込みが起きた時に発生する記憶保護割込によって呼び出される。記憶保護割込処理300は、割込要因となったメモリへの読み出し/書き込みアドレスによって(1)論理I/O デバイスに対する読み出しコマンド、もしくは、書き込みコマンドを検出し、仮想I/Oクライアント処理260 を呼び出す、(2)OSが記憶保護割込アドレスに対する書きこみをした時にOSの記憶保護割込先アドレスを登録する、(3)前記以外の理由による割込であれば、OS の記憶保護割込先アドレスに分岐する、という3つの処理を行なう。

【 0 0 2 5 】

記憶保護割込によって記憶保護割込処理300が呼び出されると、まず、記憶保護割込アドレス判定301を行ない、記憶保護割込となった原因が記憶保護割込の分岐先アドレスに対する書き込みであるかを判定する。記憶保護割込の分岐先アドレスに対する書き込みであれば、OSの記憶保護割込を登録302で、分岐先アドレスに書き込もうとしたデータを、OSの記憶保護割込へ分岐304での分岐先アドレスに登録し、記憶保護割込処理300を終了する。次に、I/Oアドレス判定303を行ない、記憶保護割込となった原因が論理I/Oデバイスのメモリアドレスに対する書き込みで、論理I/Oデバイスへの読み出しコマンドの発行、もしくは、書き込みコマンドの発行であるかを判定する。論理I/Oデバイスへの読み出しコマンドの発行、もしくは、書き込みコマンドの発行であれば、仮想I/Oクライアント処理260を呼び出し、記憶保護割込処理300を終了する。I/Oアドレスに対する書き込みでなければ、OSの記憶保護割込へ分岐304を実行し、OSの記憶保護割込に分岐する。

【 0 0 2 6 】

図7に第一、第二の実施形態で共通に用いられるサーバコンピュータ側の記憶保護割込処理380の手順の詳細を示す。記憶保護割込処理380はサーバコンピュー

タ102で実行されるハイパバイザ内のプログラムで、特定のメモリアドレスに対する読み出し/書き込みが起きた時に発生する記憶保護割込によって呼び出される。記憶保護割込処理380は、割込要因となったメモリへの読み出し/書き込みアドレスによって(1)OSがNIA受信割込アドレスに対する書き込みをした時にOSの記憶保護割込先アドレスを登録する、(2)OSが記憶保護割込アドレスに対する書き込みをした時にOSの記憶保護割込先アドレスを登録する、(3)前記以外の理由による割込であれば、OSの記憶保護割込先アドレスに分岐する、という3つの処理を行なう。

【0027】

記憶保護割込によって記憶保護割込処理380が呼び出されると、まず、記憶保護割込アドレス判定381を行ない、記憶保護割込となった原因が記憶保護割込の分岐先アドレスに対する書き込みであるかを判定する。記憶保護割込の分岐先アドレスに対する書き込みであれば、OSの記憶保護割込アドレスを登録382で、分岐先アドレスに書き込もうとしたデータを、OSの記憶保護割込へ分岐385での分岐先アドレスに登録し、記憶保護割込処理380を終了する。次に、NIA受信割込アドレス判定383を行ない、記憶保護割込となった原因がNIA受信割込の分岐先アドレスに対する書き込みであるかを判定する。NIA受信割込の分岐先アドレスに対する書き込みであれば、NIA受信割込アドレスを登録385で、分岐先アドレスに書き込もうとしたデータを、OSのNIA受信割込処理へ分岐289での分岐先アドレスに登録し、記憶保護割込処理380を終了する。NIA受信割込の分岐先アドレスに対する書き込みでなければ、OSの記憶保護割込へ分岐384を実行し、OSの記憶保護割込に分岐する。

【0028】

図8に第一の実施形態における仮想I/Oクライアント処理260の手順の詳細を示す。仮想I/Oクライアント処理は、I/Oデバイスの制御用メモリアドレスに書き込みを行なうと記憶保護割込によって呼び出され、ネットワークを介してサーバコンピュータ102に接続されているI/Oデバイス914に対して読み出し/書き込みを行なう。まず、コマンド判定261でI/Oデバイスの制御用メモリアドレスに書き込まれたデータが読み出しコマンドであるか書き込みコマンドであるかを判定する

。読み出しコマンドである場合は、読み出しコマンドの送信262で読み出しコマンドと読み出しコマンドのパラメータ(本実施形態ではセクタ番号とセクタ数)をサーバコンピュータ102 に送信を行ない、続いて読み出しデータの受信263でサーバコンピュータ102がI/O読み出し処理で得たデータを受信する。書き込みコマンドである場合は、書き込みコマンドの送信264で書き込みコマンドと書き込みコマンドのパラメータ(本実施形態ではセクタ番号とセクタ数)をサーバコンピュータ102に送信を行ない、続いて書き込みデータの送信265 で書き込みを行なうデータをサーバコンピュータ102に送信する。

【 0 0 2 9 】

図9に第一の実施形態におけるNIA受信割込処理280の手順の詳細を示す。NIA受信割込処理280はサーバコンピュータ102で実行されるハイパバイザ内のプログラムで、NIA912がネットワークからデータを受信すると割込が発生して呼び出される。NIA 受信割込処理280が呼び出されると、仮想I/Oコマンド判定281を実行し、NIAが受信したデータがクライアントコンピュータ101の読み出しコマンドの送信262、もしくは書き込みコマンドの送信264であるものかどうかを判定する。読み出しコマンドの送信262、もしくは、書き込みコマンドの送信264 である場合は、仮想I/Oサーバ処理400を実行し、それ以外の場合は、OSのNIA受信割込処理へ分岐289を実行し、OSで受信割込処理を実行する。

【 0 0 3 0 】

図10に第一の実施形態における仮想I/Oサーバ処理400の手順の詳細を示す。まず、コマンド判定288を実行し、読み出しコマンドか書き込みコマンドかを判定する。読み出しコマンドである場合は、I/Oデバイス914に対して読み出しコマンド発行282 を行ない読み出しコマンドと読み出しコマンドのパラメータ(本実施形態ではセクタ番号とセクタ数)をI/Oデバイスのメモリアドレスに書き込みを行なう。読み出しコマンド発行282 が行なわれると、I/O デバイス内でI/O読み出し処理283が実行され、I/Oデバイス914からデータの読み出しが行われる。続いて、読み出し結果の送信284 を行ない、I/Oデバイス914 から読み出したデータをクライアントコンピュータ101に送信する。書き込みコマンドである場合は、書き込みデータの受信285でクライアントコンピュータ101から書き込みデータを

受信する。書き込みコマンド発行286が行なわれると、I/O デバイス内でI/O書き込み処理287が実行され、I/Oデバイス914へデータの書き込みが行われる。

【 0 0 3 1 】

図11に第一、第二の実施形態で共通に用いられるクライアントコンピュータ101における起動時処理320の手順の詳細を示す。起動時処理320はクライアントコンピュータ101が起動する時に呼び出され、ハイパバイザの起動、OSの起動と記憶保護割込の設定が行う。

【 0 0 3 2 】

まず、プログラムローダをディスクからメモリに転送し、プログラムローダ起動321を行う。プログラムローダは、ハイパバイザをディスクからメモリに転送し、ハイパバイザ起動322を行なう。ハイパバイザは、読み出しコマンド発行213、書き込みコマンド発行363、記憶保護割込アドレスの変更を記憶保護機構を用いて監視するために、特権レベル取得323と記憶保護割込設定324を行い、I/Oデバイスのメモリアドレス、もしくは、記憶保護割込の分岐先アドレスへの書き込み、読み出しがある場合に記憶保護割込みが発生し、記憶保護割込処理300へ分岐する様に設定する。

【 0 0 3 3 】

ここで、特権レベルについて説明する。通常プロセッサはOSカーネルが動作する特権レベルと、ユーザプログラムが動作する一つ以上の非特権レベルを持っている。特権レベルでは一切の制限なく全てのハードウェアをアクセスでき、逆に非特権レベルではハードウェアへのアクセスは一切出来ないようになっている。特権レベル取得では、プロセッサの動作モードを切り替え、特権レベルを使用可能にした上で特権レベルでプログラムを動作させるようにする。

【 0 0 3 4 】

続いて、プログラムローダは、OSをディスクからメモリに転送し、OSの起動325を行う。OSは 記憶保護割込設定326 記憶保護割込の分岐アドレスにOS の記憶保護割込処理のアドレスの書き込みを行なう。この時、従来は記憶保護割込処理がOSに設定されるが、本実施例では、記憶保護割込が発生し、ハイパバイザの記憶保護割込処理300内でOSの記憶保護割込304での分岐先アドレスとして登録され

る。

【 0 0 3 5 】

図12に第一の実施形態におけるサーバコンピュータ102における起動時処理の手順の詳細を示す。起動時処理340はクライアントコンピュータ102が起動する時に呼び出され、ハイパバイザの起動、OSの起動、記憶保護割込の設定、NIA受信割込の設定を行なう。

【 0 0 3 6 】

まず、プログラムローダをディスクからメモリに転送し、プログラムローダ起動341を行う。プログラムローダは、ハイパバイザをディスクからメモリに転送し、ハイパバイザ起動342を行なう。ハイパバイザは、NIA受信割込の分岐先アドレスと記憶保護割込の分岐先アドレスの変更を記憶保護機構を用いて監視するために、特権レベル取得343と記憶保護割込設定344を行い、記憶保護割込の分岐先アドレスとNIA受信割込の分岐先アドレスへの書き込み、読み出しがある場合に記憶保護割込みが発生し、記憶保護割込処理380へ分岐する様に設定する。さらに、NIA受信割込設定345によりNIAがデータを受信した時にNIA 受信割込処理280へ分岐する様に設定する。

【 0 0 3 7 】

続いて、プログラムローダは、OSをディスクからメモリに転送し、OSの起動346を行う。OSは 記憶保護割込設定347 で記憶保護割込の分岐先アドレスにOS の記憶保護割込処理のアドレスの書き込みを行なう。この時、従来は記憶保護割込処理がOSに設定されるが、本実施例では、記憶保護割込が発生し、ハイパバイザの記憶保護割込処理380内でOSの記憶保護割込384での分岐先アドレスとして登録される。同様に、OSは NIA受信割込設定348でNIA 受信割込の分岐先アドレスにOS のNIA受信割込処理のアドレスの書き込みを行なう。この時、従来はNIA受信割込処理がOSに設定されるが、本実施例では、記憶保護割込が発生し、ハイパバイザの記憶保護割込処理380内でOSのNIA受信割込処理へ分岐289での分岐先アドレスとして登録される。以上の手順により、記憶保護割込とNIA受信割込をハイパバイザへ分岐する様にできる。

【 0 0 3 8 】

図13に第二の実施形態における仮想I/Oクライアント処理260の手順の詳細を示す。第二の実施形態では、サーバコンピュータのI/OデバイスへのコマンドをハイパバイザがサーバコンピュータのOSが解釈可能なプロトコルに変換してからコマンドをサーバコンピュータに送るところが第一の実施形態と異なる。まず、コマンド判定421でI/Oデバイスのメモリアドレスに書き込まれたデータが読み出しコマンドであるか書き込みコマンドであるかを判定する。

【0039】

読み出しコマンドである場合は、読み出しコマンドの変換422で読み出しコマンドをサーバコンピュータのOSが解釈可能なプロトコルに変換する。サーバコンピュータのOSがUNIX (X/Open Company Limitedが独占的にライセンスしている米国ならびに他の国における登録商標) であるならば、例えば、読み出しコマンドを `rshサーバコンピュータdd if=I/Oデバイス名 skip=読み出し開始ブロック番号 count=読み出しを行なうブロック数` を実行した時にサーバコンピュータに送信されるデータとする。読み出しコマンドの送信423では読み出しコマンドの変換422によって作成されたコマンドをサーバコンピュータ102 に送信する。続いて読み出しデータの受信424でサーバコンピュータ102のOSが読み出しコマンドを実行した結果を受信する。

【0040】

書き込みコマンドである場合は、書き込みコマンドの変換425で書き込みコマンドをサーバコンピュータのOSが解釈可能なプロトコルに変換する。サーバコンピュータのOSがUNIXであるならば、例えば、書き込みコマンドを `rshサーバコンピュータdd of=I/Oデバイス名 seek=書き込み開始ブロック番号 count=書き込みを行なうブロック数` を実行した時にサーバコンピュータに送信されるデータとする。書き込みコマンドの送信426は読み出しコマンドの変換425によって作成されたコマンドをサーバコンピュータ102 に送信する。続いて書き込みデータの送信427で書き込みを行なうデータをサーバコンピュータ102に送信する。

【0041】

以上のような実施の形態によって、I/Oデバイスの仮想化がハイパバイザによって行うことができるようになるため、サーバコンピュータとクライアントコン

コンピュータのオペレーティングシステムやアプリケーションプログラムを変更することなく、サーバコンピュータに接続されたI/Oデバイスであっても、クライアントコンピュータに接続されたI/Oデバイスであるかのように扱うことができる。

【0042】

また、第二の実施の形態のように、上記クライアントコンピュータのハイパバイザの仮想I/Oクライアント処理は、サーバコンピュータのオペレーティングシステムでサポートされているプロトコルでサーバコンピュータに読み出しコマンドもしくは書き込みコマンドを送信する。

【0043】

これによって、I/Oデバイスの仮想化がクライアントコンピュータのハイパバイザによって行うことができるようになるため、クライアントコンピュータへハイパバイザを導入するだけで、サーバコンピュータとクライアントコンピュータのオペレーティングシステムやアプリケーションプログラムを変更することなく、サーバコンピュータに接続されたI/Oデバイスであっても、クライアントコンピュータに接続されたI/Oデバイスであるかのように扱うことができる。

【0044】

【発明の効果】

本発明によれば、クライアントコンピュータのOSが論理I/Oデバイスにアクセスすると、ハイパバイザがネットワークを介してサーバコンピュータのI/Oデバイスへコマンドを発行する。これにより、クライアントコンピュータとサーバコンピュータのOSが異なる場合でも、OSやアプリケーションプログラムを変更することなく、クライアントコンピュータからサーバコンピュータのI/Oデバイスをアクセスすることができる。

【図面の簡単な説明】

【図1】 本発明を実施する計算機システムの構成図。

【図2】 クライアントコンピュータのハードディスクに格納されたソフトウェア。

【図3】 サーバコンピュータのハードディスクに格納されたソフトウェア。

【図4】読み出し処理を示すフローチャート。

【図5】書き込み処理を示すフローチャート。

【図6】クライアントコンピュータの記憶保護割込処理を示すフローチャート

【図7】サーバコンピュータの記憶保護割込処理を示すフローチャート。

【図8】クライアントコンピュータの仮想I/O処理を示すフローチャート。

【図9】サーバコンピュータのNIA受信割込処理を示すフローチャート。

【図10】サーバコンピュータの仮想I/O処理を示すフローチャート。

【図11】クライアントコンピュータの起動時処理のフローチャート。

【図12】サーバコンピュータの起動時処理のフローチャート。

【図13】第二の実施形態におけるクライアントコンピュータの仮想I/O処理を示すフローチャート。

【符号の説明】

210：ファイルの読み出し、211：デバイスドライバの検索、

212：デバイスドライバの呼び出し、213：読み出しコマンド発行、

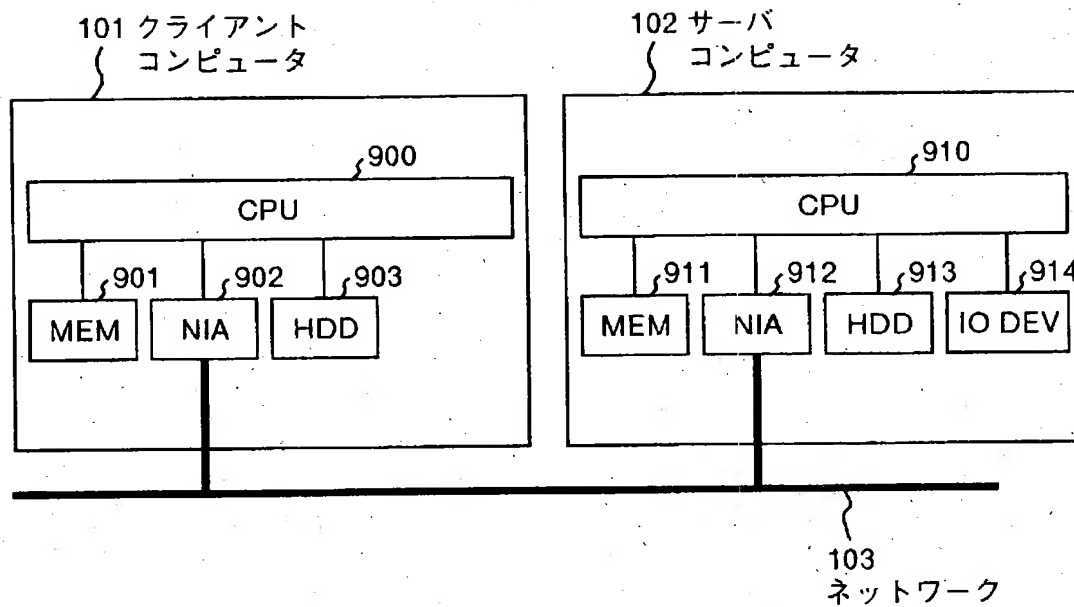
214：論理I/Oデバイスアクセス検出、

260：仮想I/Oクライアント処理。

【書類名】 図面

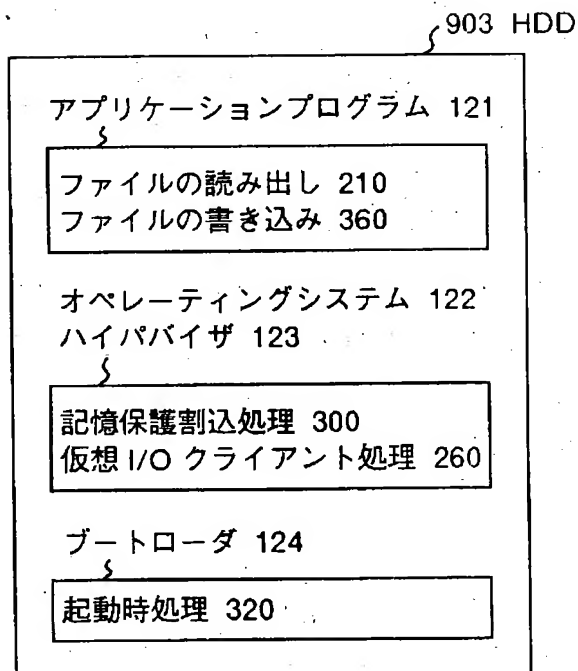
【図 1】

図 1



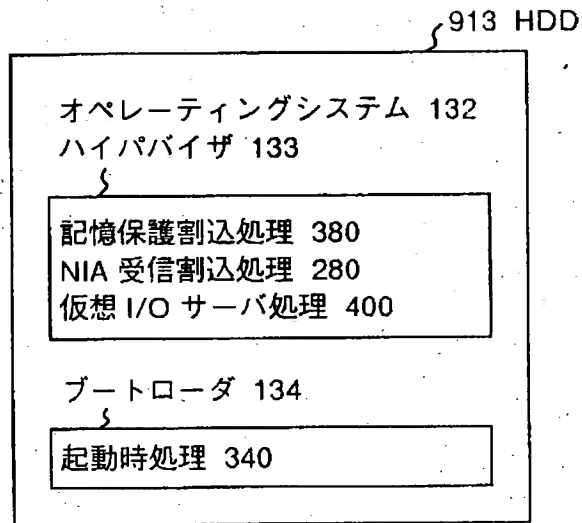
【図 2】

図 2



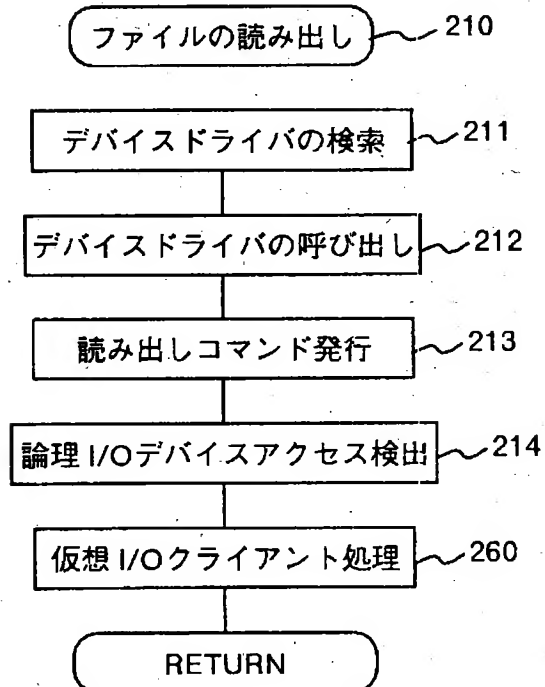
【図 3】

図 3



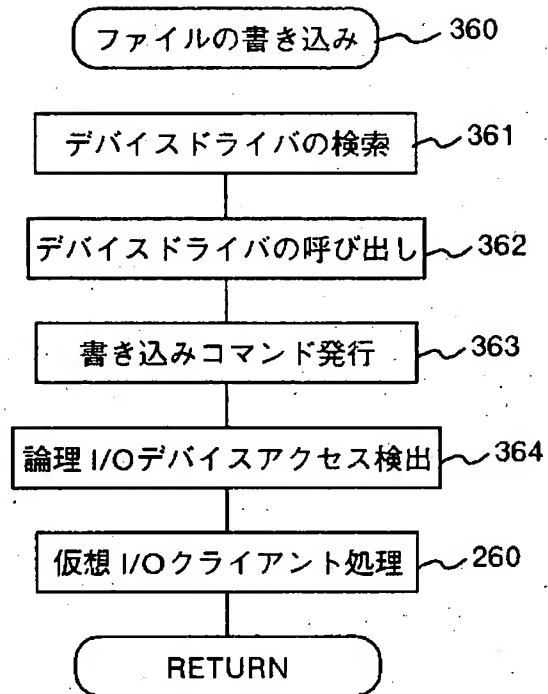
【図 4】

図 4



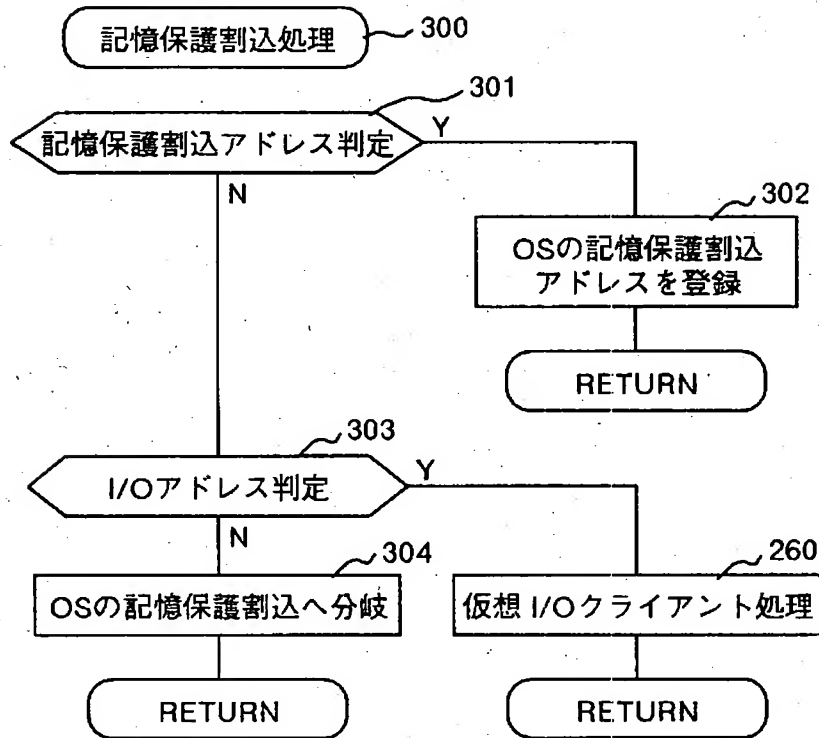
【図 5】

図 5



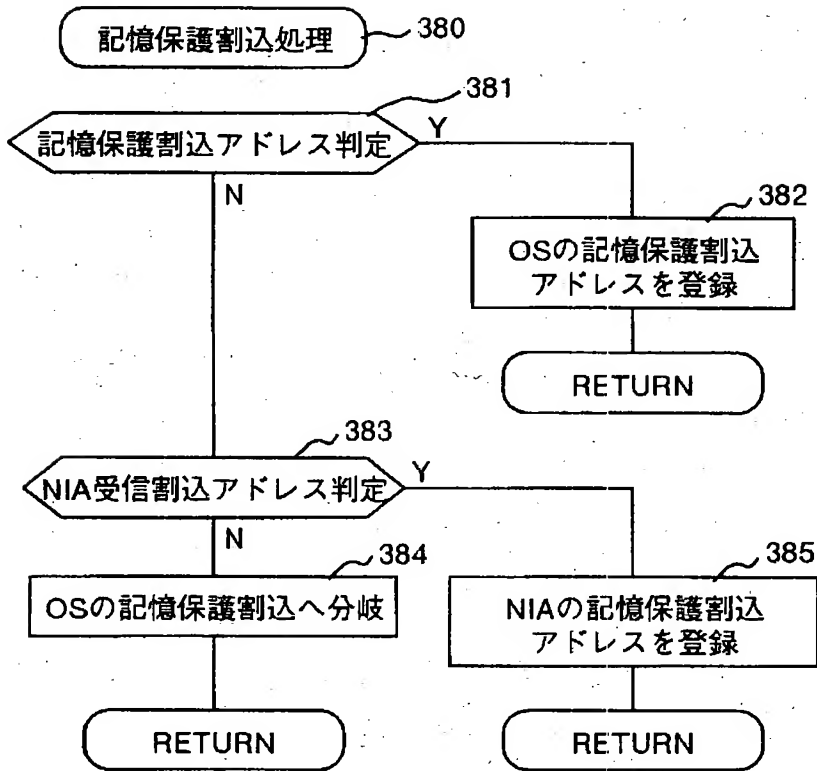
【図 6】

図 6



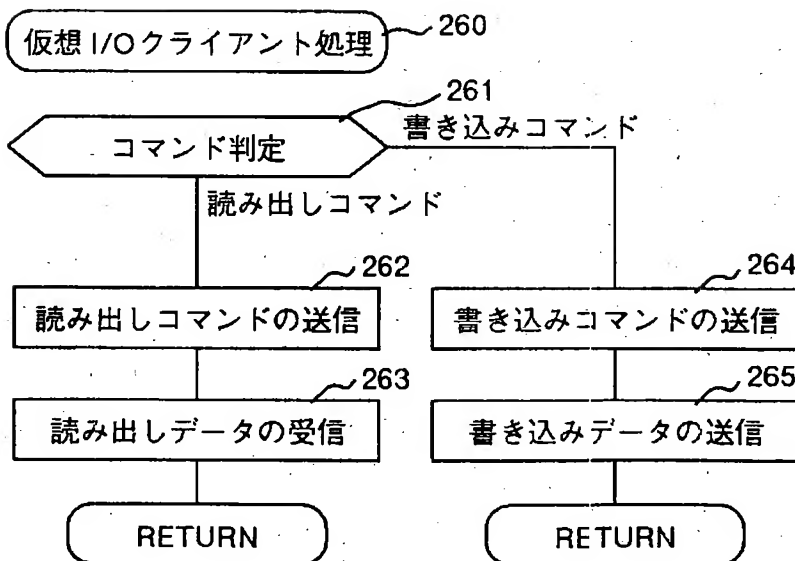
【図 7】

図 7



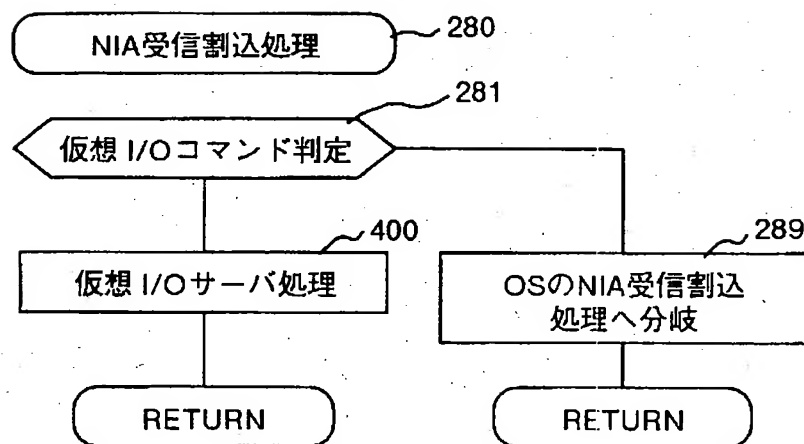
【図 8】

図 8



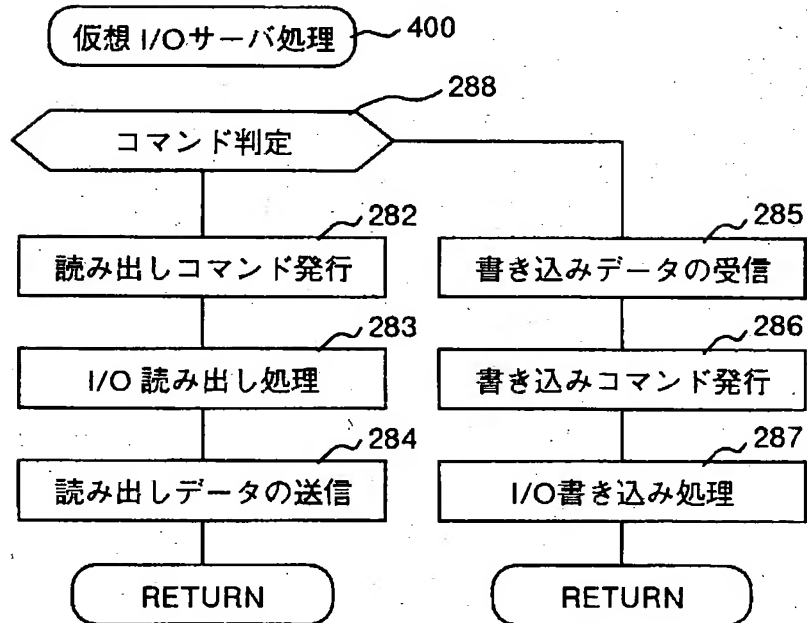
【図 9】

図 9



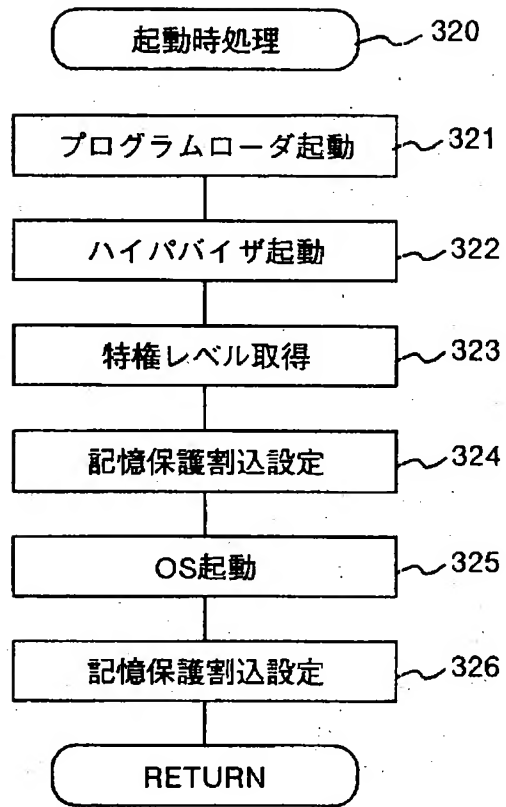
【図 10】

図 10



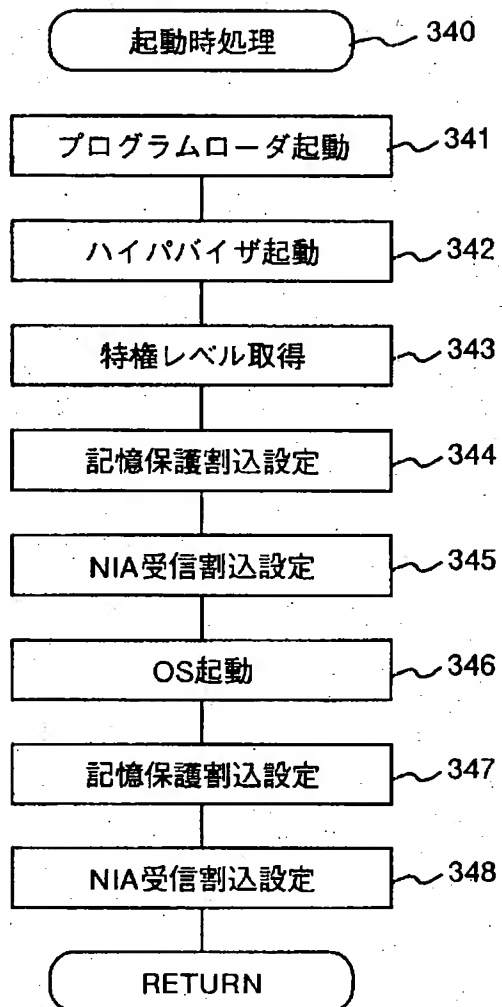
【図 1 1】

図 1 1



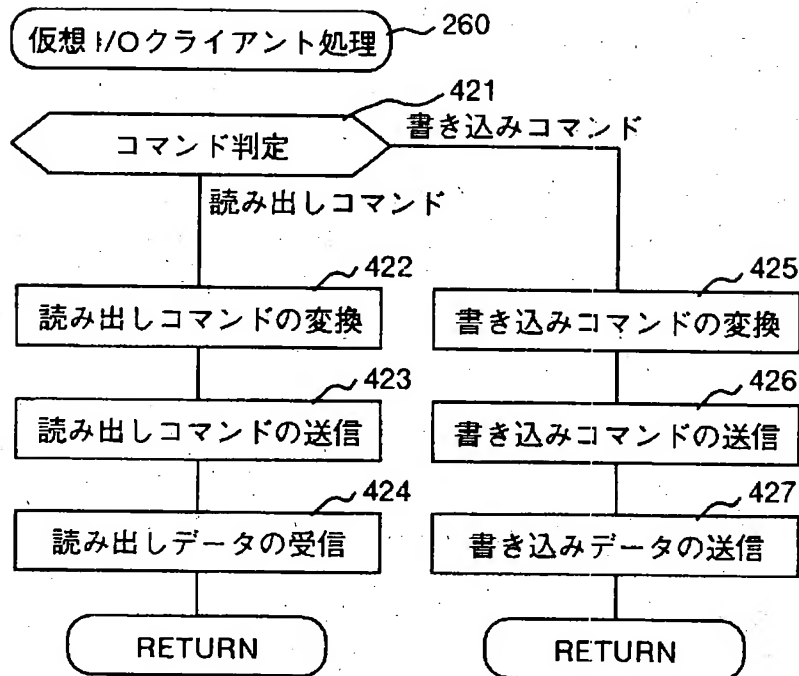
【図 1 2】

図 1 2



【図 13】

図 13



【書類名】 要約書

【要約】

【課題】 クライアントコンピュータからサーバコンピュータのI/Oデバイスを
使用可能にする。

【解決手段】 OSとハードウェアとの間に介在するハイパバイザがI/Oデバイス
の仮想化を行なう。

【選択図】 図4

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所